The Construction and Application Exploration of Medical Full-text Knowledge Extraction System based on Deep Learning

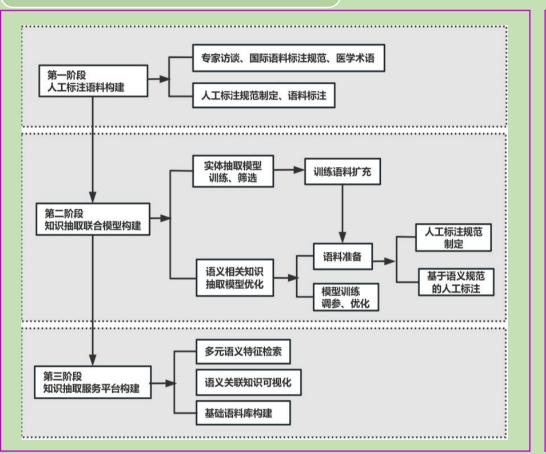


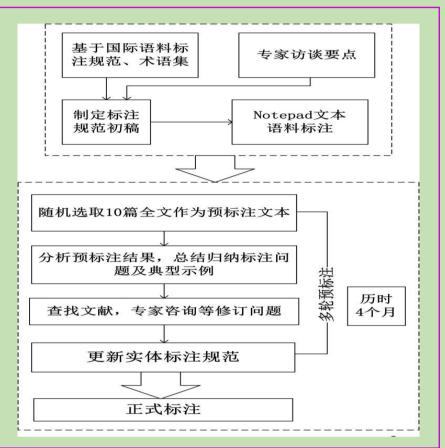
Yuan, xiaoyuan Daijuan <u>yuanxy@nju.edu.cn</u> Jiangsu, Nanjing, Nanjing University Library

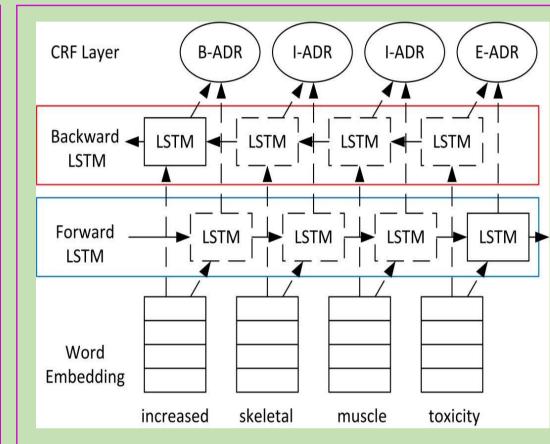
Background

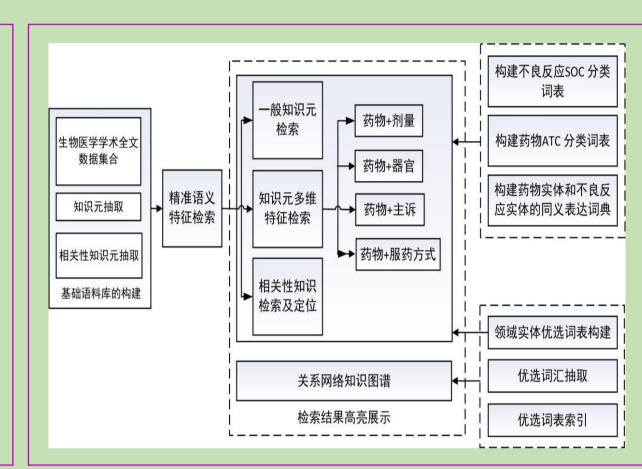
In recent years, the global volume of medical literature has experienced explosive growth. How to efficiently extract precise knowledge from massive medical literature and build intelligent knowledge service systems has remained a core challenge in the library field. This study aims to combine deep learning technologies to design efficient medical Named Entity Recognition (NER) and relationship extraction service frameworks addressing the fine-grained knowledge needs of medical researchers, while validating their effectiveness in practical application scenarios.

Work Flow



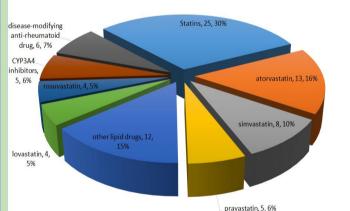






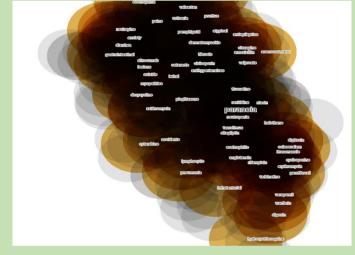
Results and Discussion

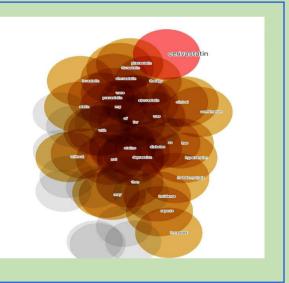












Conclusion

- .The experiment constructs a manually annotated corpus of statin-ADR relationships, followed by model training, optimization, and data augmentation.
- . The Bi-LSTM-CRF and Bi-LSTM architectures demonstrate superior performance: drug entity detection achieves F1=0.9, ADR detection F1=0.82, and Bi-LSTM outperforms attention-based At-Bi-LSTM and traditional SVM methods in drug-ADR relation extraction (F1=0.8).
- . Leveraging 1,602 annotated full-text articles, we developed a domain knowledge extraction platform (http://lk.fmcib.com.cn:19081/nlp/) featuring intelligent reading paradigms for medical researchers.

Application http://lk.fmcib.com.cn:19081/nlp/



practices or

Abstract

Background & Aims

Occasional risk of serious liver dysfunction and autoimmune hepatitis during atorvastatin therapy has been reported.

We compared the risk of hepatotoxicity in atorvastatin relative to simvastatin treatment.

Methods

The UK GPRD identified patients with a first prescription for simvastatin [164,407] or atorvastatin [76,411] between 1997 and 2006, but with no prior record of liver diske ase, alcohol-related diagnosis, or liver dysfunction.

Incident liver dysfunction in the following six months was identified by blochemical value and compared between statin groups by Cox regression model adjusting forage, sex, year treatment started, dose, alcohol consumption, smoking, body mass index and comorbid conditions.

Results

Moderate to severe hepatotoxicity (bilirubin > 60gmoV/L, AST or ALT > 200U/L or alkaline phosphatase > 1200U/L) developed in 71 patients on atorvastatin were sus 101 on simvastatin.

Adjusted hazard ratio (AHR) for all atorvastatin relative to simvastatin was 1.9 (95% confidence interval 1.4-2.6). High dose was classified as 40-80mg daily and low dose 1020mg daily.

Hepatotoxicity occurred in 0.44% of 4075 patients on high dose atorvastatin [HDA], 0.07% of 72,336 on low dose atorvastatin (LDA), 0.09% of 44,675 on high dose simvastatin [HDS] and 0.05% of 119,732 on low dose simvastatin [LDS].

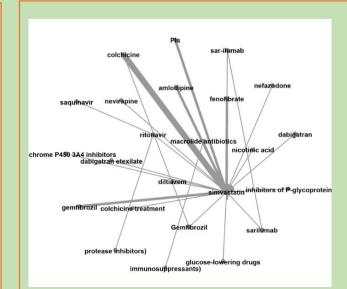
A HRS compared to LDS were 7.3 (42-12.7) for HDA, 1.4 (0.9-2.0) for LDA and 1.5 (1.0-2.2) for HDS.

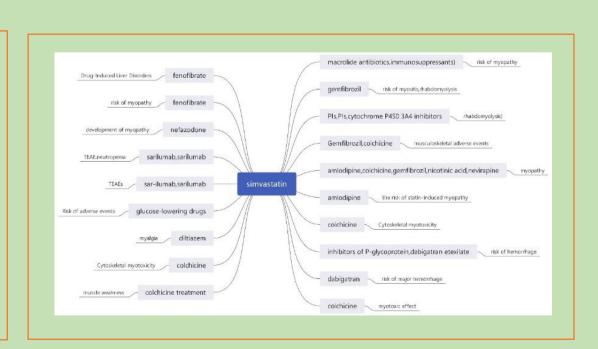
PLOS ONE [DOI:10.1371/journal.pone.0151587 March 16, 2016

1/13

Atorvastatin Hepatotoxicity in Comparison to Simvastatin in UK GPRD







Reference

Hang Wang, et al. Causal-relationship representation enhanced joint extraction model for elements and relationships[J]. Neurocomputing, 2025,613(1). Xiaocheng Luo, et al. A bi-consolidating model for joint relational triple extraction[J], Neuraocomputing, 2025,614(21):128768.

Jiandong Liu, A cascasded retrieval-while-reasoning multi-document comprehension framework with incremental attention for medical question answering[J], Expert Systems with Applications, 2025,265(15):125701.

Jinbin Zhang, More intelligent knowledge graph: A large language model-driven method for knowledge presentation in geospatial digital twins[J], International Journal of Applied Earth Observation and Geoinformation, 2025, 139:104527.